ADVANCED ANALYTICS APPLICATIONS DEVELOPMENT IN THE FINANCE SECTOR

CASE STUDY

WHEN PERFORMANCE MATTERS
The customer is a bank, made up of over twenty companies and 300 branches, which operates in multiple geographical areas in Italy and around the world, and offers a range of banking and financial products and services. In addition to having a strong experience in the asset management and in private banking, the company is a leader in innovation and research: alongside business development, the creation of new modern and convenient support tools has always been a priority. E4 has been working with this important company since 2014. Over the last year it has implemented the GAIA solution, to offer significant improvements within the development team, both for processing and calculation, and for systems engineering.
THE ISSUE

An uneven working environment between development stations and long processing times in data model training

The customer’s developer team used personal computers to develop machine learning codes, with the consequent risk of having, on one hand, a non-uniform work environment between the various stations and, on the other, long processing times in the Data Models’ training. The development of Data Driven tools to support the trader’s decisions in defining investment choices required the implementation of Data Model based on Reinforcement Learning techniques, particularly onerous from a computational point of view, but suitable for GPU Computing. The customer has therefore chosen to have a centralized system, based on GPU, to have a reserve of computing power always available for all its internal users and a coherent environment for the development team.
THE SOLUTION

TO SATISFY THESE NEEDS FULLY, E4 HAS PROPOSED:

1. A CENTRALIZED APPROACH

E4 advised the customer to use a development, training and test environment based on multi-GPU servers (specifically, two 32 GB NVIDIA TESLA V100 for PCIe), proposing the GAIA solution

www.E4Company.com/Gaia
2. A CONTAINERIZED MULTI-USER ENVIRONMENT

A centralized server-based approach with GPU power; a container-based multi-user system for the coexistence of two or more development environments, to improve the training of neural networks.

Artificial Intelligence frameworks and libraries (such as, for example, Tensorflow, Pytorch and Chainer) are constantly evolving and the release of new versions is very frequent. Without the implementation of containers, managing the coexistence of multiple versions of the same libraries is a very complex operation. The implementation of working environment in containers, on the other hand, allows both to make two or more environments coexist, and to easily update the Machine Learning and Deep Learning frameworks used by the user, in a context in which the containers always access directly to the computing resources, therefore they make optimal use of the computing resources available on the server: this is essential for the training of neural networks.

E4 has provided to the customer a multi-user environment, with a user-friendly interface (E4 Analytics Studio), based on containers. In this way, all users/developers can use the same environment without overlaps and having access to both updated and “elder” working environments at the same time. Each developer has a Jupyter Notebook interface available from its browser to work with. The development environment thus becomes a high-performance container system capable of guaranteeing the permanence of the frameworks used over time, in order to always reproduce the results already obtained in the past.

In the initial phase, the customer was skeptical towards container technology, because it was not yet thoroughly investigated, but, given the ease of management over time of the coexistence of different development environments, he agreed to accept the proposal.
3. A SIMPLE AND FLEXIBLE INTERFACE

Gaia is a highly flexible solution thanks to its development interface based on Jupyter Notebook, but also open to host more traditional development environments.

Last but not least, there is a third element, i.e. the interface provided to the internal user. E4 provided an easy-to-use interface (E4 Analytics Studio) which uses Jupyter Notebook, an environment that is the de-facto standard in Data Science. Thanks to this easy-to-use interface, the user can choose his own work environment, always having both his personal data area and an area shared between the team.

All data and codes of all users are hosted on the central system, thus making backup operations very simple and natural.
ABILITY TO INTEGRATE WITH MORE TRADITIONAL DEVELOPMENT ENVIRONMENTS

The development interface based on Jupyter Notebook is a de facto standard in the Data Science field. However, some developers still prefer to use IDE environments such as Microsoft Visual Studio. For this reason, we have given the customer the possibility to use the containerized development environments, included in GAIA, through Visual Studio Code, an open source environment, developed by a large community led by Microsoft, which reproduces the typical way of working of Visual Studio, in environments that use open source development tools.

GAIA, therefore, is an extremely flexible system, which gives the Data Scientist and Data Engineer the possibility to choose a more modern interface based on a Notebook, or to obtain a more “traditional” user experience, using the development environments containerized inside Visual Studio Code.